Best Linear Unbiased Prediction (BLUP)
Best Linear Unbiased Prediction (BLUP):

Suppose

\[ y = X\beta + \varepsilon, \]

where

\[ E(\varepsilon) = 0, \quad \text{and} \quad \text{Var}(\varepsilon) = \Sigma = \sigma^2 V. \]
Initially, we will assume the Aitken model holds so that $\sigma^2 > 0$ is unknown and $\mathbf{V}$ is a known symmetric and PD matrix.
Suppose $u$ is a random variable with mean 0 and finite variance.

A linear predictor $d + a'y$ of $c'\beta + u$ is unbiased iff

$$E(d + a'y) = E(c'\beta + u) = c'\beta \quad \forall \beta \in \mathbb{R}^p.$$
$c'\beta + u$ is predictable iff \( \exists \) an unbiased linear predictor of $c'\beta + u$. 
Note that $c'\beta + u$ is predictable $\iff \exists \ a \ni c' = a'X$.

This follows from the same argument used to show that $c'\beta$ estimable $\iff \exists \ a \ni c' = a'X$. 
Also, $d + a'y$ is an unbiased predictor of $c'\beta + u$ iff

\[ d = 0 \quad \text{and} \quad c' = a'X, \]

as argued for the case of an unbiased estimator.
Suppose $\hat{w}$ is a predictor of a random variable $w$.

The mean squared error (MSE) of $\hat{w}$ as a predictor of $w$ is

$$E(\hat{w} - w)^2.$$
Suppose

\[ \text{Cov}(\varepsilon, u) = \sigma^2 k \]

for some known vector \( k \).

Suppose \( c'\beta + u \) is predictable.

Prove that \( c'\hat{\beta}_{\text{GLS}} + \hat{u} \) has lowest MSE among all unbiased linear predictors of \( c'\beta + u \), where

\[ \hat{u} \equiv \left[ \text{Cov}(\varepsilon, u) \right]'\left[ \text{Var}(\varepsilon) \right]^{-1}(y - X\hat{\beta}_{\text{GLS}}) = k'V^{-1}(y - X\hat{\beta}_{\text{GLS}}). \]
Proof:

First note that

\[ c' \hat{\beta}_{\text{GLS}} + \hat{u} = c' \hat{\beta}_{\text{GLS}} + k' V^{-1} (y - X \hat{\beta}_{\text{GLS}}) \]

has expectation

\[ c' \beta + k' V^{-1} (X \beta - X \beta) = c' \beta. \]

Thus, \( c' \hat{\beta}_{\text{GLS}} + \hat{u} \) is an unbiased predictor of \( c' \beta + u. \)
Now note that

\[
\begin{align*}
  c'\hat{\beta}_{\text{GLS}} + \hat{u} &= c'\hat{\beta}_{\text{GLS}} + k'V^{-1}(y - X\hat{\beta}_{\text{GLS}}) \\
  &= (c' - k'V^{-1}X)\hat{\beta}_{\text{GLS}} + k'V^{-1}y \\
  &= (c' - k'V^{-1}X)(X'V^{-1}X)^{-1}X'V^{-1}y + k'V^{-1}y \\
  &= [(c' - k'V^{-1}X)(X'V^{-1}X)^{-1}X'V^{-1} + k'V^{-1}]y \\
  \equiv b'y & \quad \text{(a linear estimator.)}
\end{align*}
\]
Because
\[ c' \hat{\beta}_{GLS} + \hat{u} = b'y \]
is an unbiased predictor, we know
\[ b'X = c'. \]
Let $a'y$ be any other linear unbiased predictor of $c'\beta + u$. Then $a'X = c'$. 
The MSE of $a'y$ is

$$E[a'y - (c'\beta + u)]^2 = \text{Var}[a'y - (c'\beta + u)]$$
$$= \text{Var}(a'y - u)$$
$$= \text{Var}(a'y - b'y - u + b'y)$$
$$= \text{Var}[(a - b)'y] + \text{Var}(b'y - u)$$
$$+ 2\text{Cov}[(a - b)'y, b'y - u].$$
Now

\[
\text{Cov}[(a - b)'y, b'y - u] = (a - b)'^\text{Var}(y)b - \text{Cov}[(a - b)'y, u] \\
= (a - b)'^\sigma^2Vb - (a - b)'\text{Cov}(y, u) \\
= \sigma^2(a - b)'Vb - (a - b)'\text{Cov}(\varepsilon, u) \\
= \sigma^2(a - b)'Vb - \sigma^2(a - b)'k \\
= \sigma^2(a - b)'(Vb - k).
\]
Now

\[
Vb - k = V[V^{-1}X[(X'V^{-1}X)^{-}]]'(c - X'V^{-1}k) + V^{-1}k - k
= X[(X'V^{-1}X)^{-}][c - X'V^{-1}k].
\]

Thus, the covariance term is

\[
\sigma^2(a - b)'X[(X'V^{-1}X)^{-}][c - X'V^{-1}k],
\]

which is 0 because

\[
a'X = b'X = c'.
\]
Thus, we have

\[
\text{MSE}(a'y) = \text{Var}[(a - b)'y] + \text{Var}(b'y - u) \\
= \text{Var}[(a - b)'y] + \text{Var}(b'y - (c'\beta + u)) \\
= \text{Var}[(a - b)'y] + E[(b'y - (c'\beta + u))^2] \\
= \text{Var}[(a - b)'y] + \text{MSE}(b'y).
\]
Thus,

$$\text{MSE}(a'y) \geq \text{MSE}(b'y)$$

with equality iff

$$\text{Var}[(a - b)'y] = 0 \iff a = b.$$

Thus $c'\hat{\beta}_{\text{GLS}} + \hat{u}$ is the unique best linear unbiased predictor (BLUP) of $c'\beta + u$. □
In practice,

\[ \Sigma = \sigma^2 V \]

involves unknown variance components in addition to the unknown \( \sigma^2 \).

We replace unknown variance components in \( c' \hat{\beta}_{\text{GLS}} + \hat{u} \) by the estimates to obtain “empirical” BLUPs (eBLUPs).
This typically results in a nonlinear predictor whose properties are not so well characterized.
Example:

Suppose $y_{ij}$ is the average monthly milk production of the $j^{th}$ daughter of sire $i$.

Suppose

$$y_{ij} = \mu + s_i + e_{ij} \quad \text{for } i = 1, \ldots, t; \ j = 1, \ldots, n_i,$$

where $s_1, \ldots, s_t$ are iid with mean 0 and variance $\sigma_s^2$. 
independent of the $e_{ij}$ terms, which are iid with mean 0 and variance $\sigma_e^2$.

Suppose $\sigma_s^2/\sigma_e^2$ is a known constant $d$.

Find an expression for the BLUP of $\mu + s_1$. 
Let \( \mathbf{y} = [y_{11}, \ldots, y_{1n_1}, y_{21}, \ldots, y_{2n_2}, \ldots, y_{tn_t}]' \).

Let \( X = \mathbf{1} \) where \( n = \sum_{i=1}^{t} n_i \).

Let \( \beta = [\mu] \).

Let \( \varepsilon =
\begin{bmatrix}
s_1 + e_{11} \\
\vdots \\
s_1 + e_{1n_1} \\
s_2 + e_{21} \\
\vdots \\
s_2 + e_{2n_2} \\
\vdots \\
s_t + e_{tn_t}
\end{bmatrix}.
\)
Then $\text{Var}(\varepsilon)$ is block diagonal with blocks of sizes 
$n_1 \times n_1, n_2 \times n_2, \ldots, n_t \times n_t$.

Each block has $\sigma_s^2 + \sigma_e^2$ on the diagonal and $\sigma_s^2$ on the 
off-diagonal.

The $i^{th}$ block is 

$$
\sigma_e^2 \begin{bmatrix} I \\ n_i \times n_i \end{bmatrix} + \sigma_s^2 \begin{bmatrix} 11' \\ n_i \times n_i \end{bmatrix} = \sigma_e^2 \left[ I + d11' \right].
$$
We wish to predict \( c'\beta + u \), where \( c = [1] \), \( \beta = [\mu] \) and \( u = s_1 \).

\[
\text{Cov}(\varepsilon, u) = \text{Cov}(\varepsilon, s_1) = \begin{bmatrix} \sigma^2_s & 1 \\ n_1 \times 1 & 0 \\ (n-n_1) \times 1 \end{bmatrix}
\]

\[
= \sigma^2_e \begin{bmatrix} d & 1 \\ n_1 \times 1 & 0 \\ (n-n_1) \times 1 \end{bmatrix} \equiv \sigma^2_e k.
\]
We need $\hat{\beta}_{\text{GLS}}$, a solution to the Aitken equations

$$X'V^{-1}X\beta = X'V^{-1}y,$$

where

$$V = \text{diag}(I_{n_1 \times n_1} + d_{11'}, \ldots, I_{n_t \times n_t} + d_{11'}).$$

$$\left(I_{n_i \times n_i} + d_{11'}^{n_i \times n_i}\right)^{-1} = I_{n_i \times n_i} - \frac{d}{1 + dn_i^{n_i \times n_i} 11'}. $$
Thus,

\[
X'V^{-1}X = 1'V^{-1}1 \\
= \sum_{i=1}^{t} 1' \left[ I - \frac{d}{1 + dn_i} 11' \right] 1 \\
= \sum_{i=1}^{t} \left( n_i - \frac{dn_i^2}{1 + dn_i} \right) \\
= \sum_{i=1}^{t} \frac{n_i}{1 + dn_i}.
\]
\[ X'V^{-1}y = \sum_{i=1}^{t} 1' \left[ I - \frac{d}{1 + dn_i} 11' \right] y_i \]

(where \( y_i = [y_{i1}, \ldots, y_{in_i}]' \))

\[ = \sum_{i=1}^{t} \left( y_i' - \frac{dn_i}{1 + dn_i} y_i' \right) \]

\[ = \sum_{i=1}^{t} \left( 1 - \frac{dn_i}{1 + dn_i} \right) y_i' \]

\[ = \sum_{i=1}^{t} \frac{n_i}{1 + dn_i} \bar{y}_i' \]
Thus,

\[
\hat{\beta}_{\text{GLS}} = \hat{\mu}_{\text{GLS}} = (X'V^{-1}X)^{-1}X'V^{-1}y \\
= \frac{\sum_{i=1}^{t} \frac{n_i}{1+dn_i} \bar{y}_i}{\sum_{i=1}^{t} \frac{n_i}{1+dn_i}} \\
= \frac{\sum_{i=1}^{t} \left( \frac{\sigma_e^2}{n_i} + \sigma_s^2 \right)^{-1} \bar{y}_i}{\sum_{i=1}^{t} \left( \frac{\sigma_e^2}{n_i} + \sigma_s^2 \right)^{-1}}.
\]
Note that each $\bar{y}_i$ is an unbiased estimator of $\mu$ with variance

$$\frac{\sigma^2_e}{n_i} + \sigma^2_s.$$ 

Thus, $\hat{\mu}_{GLS}$ is a weighted average of independent unbiased estimators of $\mu$, with inverse variances of estimators as weights.
Now,

\[
\hat{s}_1 = \hat{u} = k'V^{-1}(y - X\hat{\beta}_{\text{GLS}})
\]

\[
k'V^{-1}y = d1' \left[ I - \frac{d}{1 + dn_1}11' \right] y_1
\]

\[
= dy_1. - \frac{d^2n_1}{1 + dn_1}y_1.
\]

\[
= \frac{d + d^2n_1 - d^2n_1}{1 + dn_1}y_1.
\]

\[
= \frac{dn_1}{1 + dn_1} \bar{y}_1.
\]
\[ k'V^{-1}X\hat{\beta}_{GLS} = d1' \left[ I - \frac{d}{1 + dn_1}11' \right] 1\hat{\mu}_{GLS} \]
\[ = \left( dn_1 - \frac{d^2 n_1^2}{1 + dn_1} \right) \hat{\mu}_{GLS} \]
\[ = \frac{dn_1}{1 + dn_1} \hat{\mu}_{GLS}. \]
Thus, the BLUP of $\mu + s_1$ is

$$
\hat{\mu}_{GLS} + \frac{dn_1}{1 + dn_1} \bar{y}_1. - \frac{dn_1}{1 + dn_1} \hat{\mu}_{GLS}
$$

$$
= \frac{dn_1}{1 + dn_1} \bar{y}_1. + \left(1 - \frac{dn_1}{1 + dn_1}\right) \hat{\mu}_{GLS}
$$

$$
= \frac{\sigma_s^2}{\sigma_e^2/n_1 + \sigma_s^2} \bar{y}_1. + \frac{\sigma_e^2/n_1}{\sigma_e^2/n_1 + \sigma_s^2} \hat{\mu}_{GLS}.
$$
The mean squared error (MSE) of a predictor \( \hat{w} \) of a random vector \( w \) is

\[
\text{MSE}(\hat{w}) = E[(\hat{w} - w)(\hat{w} - w)'].
\]
Now suppose $u$ is a random vector with mean $0$ and 
\[ \text{Cov}(\varepsilon, u) = \sigma^2 K. \]

The BLUP of predictable $C\beta + u$ is $C\hat{\beta}_{\text{GLS}} + \hat{u}$, where

\[ \hat{u} \equiv [\text{Cov}(\varepsilon, u)]' [\text{Var}(\varepsilon)]^{-1} (y - X\hat{\beta}_{\text{GLS}}) = K'V^{-1} (y - X\hat{\beta}_{\text{GLS}}). \]
\( C\hat{\beta}_{\text{GLS}} + \hat{u} \) is the “best” linear unbiased predictor in the sense that

\[
\text{MSE}(\hat{w}) - \text{MSE}(C\hat{\beta}_{\text{GLS}} + \hat{u})
\]

is nonnegative definite \( \forall \hat{w} \), a linear unbiased predictor of \( C\beta + u \).
As for the case of univariate prediction, we replace unknown variance components in $C\hat{\beta}_{\text{GLS}} + \hat{u}$ by their estimates to obtain an eBLUP whenever necessary.
For example, suppose

\[ y = X\beta + Zu + e, \quad \text{where} \]

\[ E(u) = 0 \quad E(e) = 0 \]

\[ \text{Var}(u) = G \quad \text{Var}(e) = R \]

\[ \text{Cov}(u, e) = 0. \]
Then, with $\varepsilon = Zu + e$, we have

$$\text{Var}(\varepsilon) = ZGZ' + R$$

and

$$\text{Cov}(\varepsilon, u) = \text{Cov}(Zu + e, u)$$
$$= \text{Cov}(Zu, u)$$
$$= Z\text{Cov}(u, u)$$
$$= Z\text{Var}(u)$$
$$= ZG.$$
It follows that the BLUP of $u$ is

$$GZ'(ZGZ' + R)^{-1}(y - X\hat{\beta}_{GLS}).$$
Similarly, if $D$ is a known $m \times q$ matrix, the BLUP of $Du$ is

$$DGZ'(ZGZ' + R)^{-1}(y - X\hat{\beta}_{GLS}).$$
When $G$ and $R$ are unknown, we use the eBLUP

$$D\hat{G}Z'(\hat{Z}GZ' + \hat{R})^{-1}(y - X\hat{\beta}_{GLS}).$$
Note that computing $\hat{\beta}_{GLS}$ or computing the BLUP of $Du$ requires inversion of the $n \times n$ matrix $ZGZ' + R$.

This can be computationally expensive.
We often assume that

\[ R = \sigma_e^2 I, \]

but even then

\[ ZGZ' + \sigma_e^2 I \]

is a \( n \times n \) matrix that may be difficult to invert.
Until further notice, assume

\[ R = \sigma_e^2 I \]

for some unknown \( \sigma_e^2 > 0 \).

Define \( H = \frac{1}{\sigma_e^2} G \). Then

\[
\text{Var}(y) = ZGZ' + \sigma_e^2 I
\]

\[
= \sigma_e^2 (ZHZ' + I) = \sigma_e^2 V
\]

where \( V = ZHZ' + I \).
C.R. Henderson introduced the Mixed Model Equations (MME)

\[
\begin{bmatrix} X'X & X'Z \\ Z'X & H^{-1} + Z'Z \end{bmatrix} \begin{bmatrix} \tilde{\beta} \\ \tilde{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}.
\]
Recall that the Aitken equations are

\[ X'V^{-1}X\tilde{\beta} = X'V^{-1}y, \]

which in our special case are equivalent to

\[ X'(ZHZ' + I)^{-1}X\tilde{\beta} = X'(ZHZ' + I)^{-1}y. \]
Henderson showed that

1. The MME are consistent.

2. If \( \begin{bmatrix} \tilde{\beta} \\ \tilde{u} \end{bmatrix} \) solve the MME, then \( \tilde{\beta} \) is a solution to the AE (equivalently, \( X\tilde{\beta} = X\hat{\beta}_{\text{GLS}} \)) and \( \tilde{u} \) is the BLUP of \( u \).

3. Conversely, if \( \tilde{\beta} \) solves the AE, then the MME have a solution whose leading subvector is \( \tilde{\beta} \).
A nice thing about this result is that we can find $\hat{\beta}_{\text{GLS}}$ and the BLUP of $u$ (or $Du$) without inverting the $n \times n$ matrix $(ZGZ' + \sigma_e^2 I)$.

$$
\begin{bmatrix}
X'X & X'Z \\
Z'X & H^{-1} + Z'Z
\end{bmatrix}
$$

is $(p + q) \times (p + q)$. In some problems, $p + q \ll n$. 
By (2), the BLUP of a predictable $C\beta + Du$ is given by

$$C\tilde{\beta} + D\tilde{u}$$

where $\begin{bmatrix} \tilde{\beta} \\ \tilde{u} \end{bmatrix}$ solves the MME.
To prove Henderson’s results, we begin with the following lemma.

**Lemma H1:** Suppose $A$ is a symmetric and positive definite matrix. Suppose $W$ is any matrix with number of columns equal to number of rows of $A$. Then

$$AW'(I + WAW')^{-1} = (A^{-1} + W'W)^{-1}W'.$$
Proof:

First show that

\[ I + WAW' \quad \text{and} \quad A^{-1} + WW' \]

are each nonsingular.
\[ \forall x \neq 0, \; x'(I + WAW')x = x'x + x'WAW'x \]
\[ = \|x\|^2 + \|A^{1/2}W'x\|^2 \]
\[ \geq \|x\|^2 > 0. \]

Thus, \( I + WAW' \) is PD and \( \therefore \) nonsingular.

\[ \forall x \neq 0, \; x'(A^{-1} + W'W)x = x'A^{-1}x + x'W'Wx \]
\[ = x'A^{-1}x + \|Wx\|^2 \]
\[ \geq x'A^{-1}x > 0 \]

\( \therefore A \) PD \( \implies A^{-1} \) PD. \( \therefore A^{-1} + W'W \) is nonsingular.
Now show that

\[(A^{-1} + W'W)AW' = W'(I + WAW')\]

and then complete the proof.
\[(A^{-1} + WW')AW' = A^{-1}AW' + WW'WAW'\]
\[= W' + WW'WAW'\]
\[= W'(I + WAW').\]
Multiplying on the left by

$$(A^{-1} + W'W)^{-1}$$

gives

$$AW' = (A^{-1} + W'W)^{-1}W'(I + WAW').$$

Multiplying on the right by $(I + WAW')^{-1}$ gives the result. □
Now use Lemma H1 to prove Lemma H2:

\[ V^{-1} = (I + ZHZ')^{-1} \]

\[ = I - Z(H^{-1} + Z'Z)^{-1}Z' \]
Proof:

Applying the Lemma H1 with $H = A$ and $Z = W$ yields

$$(H^{-1} + Z'Z)^{-1}Z' = HZ'(I + ZHZ')^{-1}. \quad (*)$$

Thus,

$$[I - Z(H^{-1} + Z'Z)^{-1}Z'](I + ZHZ')$$

$$= [I - ZHZ'(I + ZHZ')^{-1}](I + ZHZ') \quad \text{by (⁎)}$$

$$= I + ZHZ' - ZHZ' = I.$$
Now let’s prove Henderson’s results beginning with (2).

Suppose \( \begin{bmatrix} \tilde{\beta} \\ \tilde{u} \end{bmatrix} \) solves the MME.

Then

\[
X'X\tilde{\beta} + X'Z\tilde{u} = X'y
\]

and

\[
Z'X\tilde{\beta} + (H^{-1} + Z'Z)\tilde{u} = Z'y.
\]
The 2nd equation

\[ (H^{-1} + Z'Z)\tilde{u} = Z'y - Z'X\tilde{\beta} \]

\[ \tilde{u} = (H^{-1} + Z'Z)^{-1}Z'(y - X\tilde{\beta}) \]

\[ \tilde{u} = HZ'(I + ZHZ')^{-1}(y - X\tilde{\beta}) \text{ by } (*) \]

\[ \tilde{u} = HZ'V^{-1}(y - X\tilde{\beta}) = GZ'\Sigma^{-1}(y - X\tilde{\beta}). \]
Now this last expression for $\tilde{u}$ indicates that $\tilde{u}$ is the BLUP for $u$ as long as $\tilde{\beta}$ solves the Aitken equations.

To show this, we examine the $1^{st}$ MME.
The 1\textsuperscript{st} MME is
\[ X'X\hat{\beta} + X'Z\hat{u} = X'y. \]

We have shown
\[ \hat{u} = (H^{-1} + Z'Z)^{-1}Z'(y - X\hat{\beta}). \]

Combining these equations gives
\[ X'X\hat{\beta} + X'Z(H^{-1} + Z'Z)^{-1}Z'(y - X\hat{\beta}) = X'y. \]
By Lemma H2,

\[ Z(H^{-1} + Z'Z)^{-1}Z' = I - V^{-1}. \]

Thus,

\[ X'X\tilde{\beta} + X'(I - V^{-1})(y - X\tilde{\beta}) = X'y \]
\[ \iff X'X\tilde{\beta} + X'y - X'X\tilde{\beta} - X'V^{-1}y + X'V^{-1}X\tilde{\beta} = X'y \]
\[ \iff X'V^{-1}X\tilde{\beta} = X'V^{-1}y. \]

∴ \( \tilde{\beta} \) solves the AE and Henderson's result (2) follows.
Now let’s prove (3).

Suppose $\tilde{\beta}$ solves the AE.

Let

$$
\tilde{u} = HZ'V^{-1}(y - X\tilde{\beta})
$$

$$
= HZ'(I + ZHZ')^{-1}(y - X\tilde{\beta})
$$

$$
= (H^{-1} + Z'Z)^{-1}Z'(y - X\tilde{\beta}) \text{ by (\star)}
$$

$$
\implies (H^{-1} + Z'Z)\tilde{u} = Z'(y - X\tilde{\beta}) = Z'y - Z'X\tilde{\beta}
$$

$$
\implies Z'X\tilde{\beta} + (H^{-1} + Z'Z)\tilde{u} = Z'y.
$$
∴ 2^{nd} MME is satisfied.

Now

\[ X'V^{-1}X\tilde{\beta} = X'V^{-1}y \]

\[ \implies X'[I - Z(H^{-1} + Z'Z)Z']X\tilde{\beta} = X'[I - Z(H^{-1} + Z'Z)Z']y \]

(By Lemma H2)
It follows that

\[ X'X\tilde{\beta} + X'Z(H^{-1} + Z'Z)^{-1}Z'(y - X\tilde{\beta}) = X'y. \]

Now

\[ \tilde{u} = (H^{-1} + Z'Z)^{-1}Z'(y - X\tilde{\beta}) \]

\[ \implies X'X\tilde{\beta} + X'Z\tilde{u} = X'y. \]

\[ \therefore \text{The } 1^{st} \text{ MME is satisfied and result (3) follows.} \]
Result (1) is now easy to prove as follows.

The AE are consistent.

Thus, by Henderson’s result (3), the MME are consistent. □
Now suppose $y = X\beta + Zu + e$, where all is as before except that, instead of $\text{Var}(e) = \sigma_e^2 I$, $\text{Var}(e) = R$, a symmetric, positive definite variance matrix.

Let $S = \frac{1}{\sigma_e^2} R$, where $\sigma_e^2$ is a positive variance parameter.

Consider the transformation

$$S^{-1/2}y = S^{-1/2}X\beta + S^{-1/2}Zu + S^{-1/2}e.$$
Then \( S^{-1/2}e \sim N(0, \sigma_e^2 I) \), and our previous results with

\[
S^{-1/2}y, S^{-1/2}X, \text{ and } S^{-1/2}Z \text{ in place of } y, X, \text{ and } Z
\]

yield the mixed model equations

\[
\begin{bmatrix}
X'S^{-1}X & X'S^{-1}Z \\
Z'S^{-1}X & H^{-1} + Z'S^{-1}Z
\end{bmatrix}
\begin{bmatrix}
\tilde{\beta} \\
\tilde{u}
\end{bmatrix} =
\begin{bmatrix}
X'S^{-1}y \\
Z'S^{-1}y
\end{bmatrix}.
\]
Multiplying both sides by \(1/\sigma_e^2\) gives

\[
\begin{bmatrix}
X'R^{-1}X & X'R^{-1}Z \\
Z'R^{-1}X & G^{-1} + Z'R^{-1}Z
\end{bmatrix}
\begin{bmatrix}
\tilde{\beta} \\
\tilde{u}
\end{bmatrix}
= \begin{bmatrix}
X'R^{-1}y \\
Z'R^{-1}y
\end{bmatrix}.
\]

This is the more general form of Henderson’s Mixed Model Equations.
Finally, let $\hat{G}$ and $\hat{R}$ denote $G$ and $R$ with unknown parameters replaced by their REML estimates.

In practice, we solve

$$
\begin{bmatrix}
X'\hat{R}^{-1}X & X'\hat{R}^{-1}Z \\
Z'\hat{R}^{-1}X & \hat{G}^{-1} + Z'\hat{R}^{-1}Z
\end{bmatrix}
\begin{bmatrix}
\tilde{\beta} \\
\tilde{u}
\end{bmatrix}
= 
\begin{bmatrix}
X'\hat{R}^{-1}y \\
Z'\hat{R}^{-1}y
\end{bmatrix},
$$

to get solutions needed for computation of BLUEs and BLUPs.